Clinical Cancer Research

AACR

A New Molecular Predictor of Distant Recurrence in ER-Positive, HER2-Negative Breast Cancer Adds Independent Information to Conventional Clinical Risk Factors

Martin Filipits, Margaretha Rudas, Raimund Jakesz, et al.

Updated Version	Access the most recent version of this article at: doi:10.1158/1078-0432.CCR-11-0926
Supplementary Material	Access the most recent supplemental material at: http://clincancerres.aacrjournals.org/content/suppl/2011/08/01/1078-0432.CCR-11-0926.DC1.html

Cited Articles	This article cites 17 articles, 8 of which you can access for free at: http://clincancerres.aacrjournals.org/content/17/18/6012.full.html#ref-list-1

E-mail alerts	Sign up to receive free email-alerts related to this article or journal.
Reprints and Subscriptions	To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.
Permissions	To request permission to re-use all or part of this article, contact the AACR Publications Department at permissions@aacr.org.

A New Molecular Predictor of Distant Recurrence in ER-Positive, HER2-Negative Breast Cancer Adds Independent Information to Conventional Clinical Risk Factors

Martin Filipits[1], Margaretha Rudas[2], Raimund Jakesz[3], Peter Dubsky[3], Florian Fitzal[3], Christian F. Singer[4], Otto Dietze[7], Richard Greil[8], Andrea Jelen[5], Paul Sevelda[6], Christa Freibauer[9], Volkmar Müller[10], Fritz Jänicke[10], Marcus Schmidt[11], Heinz Kölbl[11], Achim Rody[12], Manfred Kaufmann[12], Werner Schroth[13], Hiltrud Brauch[13], Matthias Schwab[13], Peter Fritz[13,14], Karsten E. Weber[16], Inke S. Feder[15], Guido Hennig[15], Ralf Kronenwett[16], Mathias Gehrmann[15], and Michael Gnant[3], for the EP Investigators

Abstract

Purpose: According to current guidelines, molecular tests predicting the outcome of breast cancer patients can be used to assist in making treatment decisions after consideration of conventional markers. We developed and validated a gene expression signature predicting the likelihood of distant recurrence in patients with estrogen receptor (ER)–positive, HER2-negative breast cancer treated with adjuvant endocrine therapy.

Experimental Design: RNA levels assessed by quantitative reverse transcriptase PCR in formalin-fixed, paraffin-embedded tumor tissue were used to calculate a risk score (Endopredict, EP) consisting of eight cancer-related and three reference genes. EP was combined with nodal status and tumor size into a comprehensive risk score, EPclin. Both prespecified risk scores including cutoff values to determine a risk group for each patient (low and high) were validated independently in patients from two large randomized phase III trials [Austrian Breast and Colorectal Cancer Study Group (ABCSG)-6: $n = 378$, ABCSG-8: $n = 1,324$].

Results: In both validation cohorts, continuous EP was an independent predictor of distant recurrence in multivariate analysis (ABCSG-6: $P = 0.010$, ABCSG-8: $P < 0.001$). Combining Adjuvant!Online, quantitative ER, Ki67, and treatment with EP yielded a prognostic power significantly superior to the clinicopathologic factors alone [c-indices: 0.764 vs. 0.750, $P = 0.024$ (ABCSG-6) and 0.726 vs. 0.701, $P = 0.003$ (ABCSG-8)]. EPclin had c-indices of 0.788 and 0.732 and resulted in 10-year distant recurrence rates of 4% and 4% in EPclin low-risk and 28% and 22% in EPclin high-risk patients in ABCSG-6 ($P < 0.001$) and ABCSG-8 ($P < 0.001$), respectively.

Conclusions: The multigene EP risk score provided additional prognostic information to the risk of distant recurrence of breast cancer patients, independent from clinicopathologic parameters. The EPclin score outperformed all conventional clinicopathologic risk factors. *Clin Cancer Res; 17(18); 6012–20. ©2011 AACR.*

Introduction

Current guidelines for the treatment of early-stage breast cancer recommend adjuvant chemotherapy for most patients with estrogen receptor (ER)-negative or HER2-positive tumors (1, 2). In contrast, selecting chemotherapy for patients with ER-positive, HER2-negative disease is more challenging as these patients represent a wide

Authors' Affiliations: Departments of [1]Medicine I, [2]Pathology, [3]Surgery, and [4]Special Gynecology, Medical University of Vienna; [5]Department of Pathology and Microbiology, Hanusch Hospital; [6]Department of Gynecology, Hietzing Hospital, Vienna; Departments of [7]Pathology and [8]Internal Medicine III, Paracelsus Private Medical University, Salzburg; [9]Department of Pathology, Weinviertel Hospital, Mistelbach, Austria; [10]Department of Gynecology, University Medical Center Hamburg-Eppendorf, Hamburg; [11]Department of Gynecology and Obstetrics, University of Mainz, Mainz; [12]Department of Obstetrics and Gynecology, J.W. Goethe University, Frankfurt; [13]Dr. Margarete Fischer-Bosch Institute of Clinical Pharmacology Stuttgart and University Tübingen; [14]Department of Clinical Pathology, Robert Bosch Hospital, Stuttgart; [15]Siemens Healthcare Diagnostics; and [16]Sividon Diagnostics, Cologne, Germany

Corresponding Author: Martin Filipits, Institute of Cancer Research, Department of Medicine I, Medical University of Vienna, Borschkegasse 8a, A-1090 Vienna, Austria. Phone: 43-1-4277-65237; Fax: 43-1-4277-65196; E-mail: martin.filipits@meduniwien.ac.at

American Association for *Cancer Research*

Translational Relevance

According to current guidelines, molecular tests predicting the outcome of breast cancer patients can be used to assist in making treatment decisions after consideration of conventional markers. In this study, we developed and validated a gene expression signature predicting the likelihood of distant recurrence in patients with estrogen receptor–positive, HER2-negative breast cancer treated with adjuvant endocrine therapy only. The risk score (Endopredict, EP) provided additional prognostic information to the risk of distant recurrence of breast cancer patients, independent from clinicopathologic parameters, in particular Adjuvant!Online and Ki67 labeling index. By combining the EP risk score with nodal status and tumor size, we identified breast cancer patients with very low risk for distant recurrence who may be adequately and sufficiently treated with adjuvant endocrine therapy only. This new molecular test has a strong potential to assist in optimizing adjuvant therapy and thus might improve management of patients with early-stage breast cancer.

spectrum of different risk profiles: women who derive little benefit from chemotherapy in addition to endocrine therapy and women with high-risk disease where chemotherapy is very appropriate. Current guidelines recommend the use of validated multigene tests for the decision whether to add chemotherapy or not after consideration of conventional markers (1). Recently, several molecular tests for prediction of breast cancer prognosis have been developed to aid clinical decision making (3–5). Some tests have been developed in heterogeneous patient cohorts (6, 7). Therefore, a clear answer about their prognostic or predictive value in ER-positive and HER2-negative patients remains elusive because all published validation results are based on cohorts that include HER2-positive patients. Moreover, none of the tests published to date have been shown to outperform a combination of hormone receptors, HER2 and Ki67 assayed by immunohistochemistry (IHC).

Here, we present the validation of Endopredict (EP), a new RNA-based multigene score predicting the likelihood of distant recurrence in patients with early-stage, ER-positive, and HER2-negative breast cancer treated with adjuvant endocrine therapy only. The EP risk score and its combination with the clinical risk factors, tumor size and nodal status (EPclin), were generated in a large training set of 964 tumor samples. The prespecified risk scores and cutoff values were then validated in 2 independent external validation cohorts of 1,702 formalin-fixed, paraffin-embedded (FFPE) samples. The patients who donated the samples had been enrolled in 2 large randomized trials of the Austrian Breast and Colorectal Cancer Study Group (ABCSG) and were treated with adjuvant endocrine therapy only. The aim of the study was to show that the EP score can predict distant

recurrence in both cohorts and can significantly add information to standard clinicopathologic factors.

Materials and Methods

EP and EPclin risk scores

A detailed description of the training process of EP and EPclin risk scores including sample cohorts, selection of the final set of genes, algorithm generation, and threshold definition is shown in the Supplementary Appendix.

In brief, the final training set for the development of the EP and EPclin risk score consisted of 964 ER-positive, HER2-negative tumors from patients treated with adjuvant tamoxifen only. Because the IHC-based hormone receptor status was not available for all samples, selection was based on the *ESR1*/*ERBB2* gene expression using prespecified cutoff levels (Supplementary Appendix, section 2.4). In a top-down approach, we developed and defined the EP score consisting of 8 cancer-related genes of interest (GOI: *BIRC5*, *UBE2C*, *DHCR7*, *RBBP8*, *IL6ST*, *AZGP1*, *MGP*, and *STC2*) and 3 normalization genes (*CALM2*, *OAZ1*, and *RPL37A*; Supplementary Tables S1 and 4).

The design and calculation of the final prespecified EP score as used for validation is described as follows: Mathematically, EP is arranged as a linear combination to predict distant recurrence (Supplementary Fig. S4). Relative expression of each GOI was assessed as delta cycle threshold (ΔC_t) values based on normalization on the average of 3 reference genes (*CALM2*, *OAZ1*, and *RPL37A*):

$$\Delta C_t(\text{GOI}) = 20 - C_t(\text{GOI}) + [C_t(CALM2) + C_t(OAZ1) + C_t(RPL37A)]/3 \quad \text{(A)}$$

The ΔC_t values were combined into the predictive unscaled risk score s_u.

$$s_u = 0.41 \cdot \Delta C_t(BIRC5) - 0.35 \cdot \Delta C_t(RBBP8) \\ + 0.39 \cdot \Delta C_t(UBE2C) - 0.31 \cdot \Delta C_t(IL6ST) \\ - 0.26 \cdot \Delta C_t(AZGP1) + 0.39 \cdot \Delta C_t(DHCR7) \\ - 0.18 \cdot \Delta C_t(MGP) - 0.15 \cdot \Delta C_t(STC2) - 2.63 \quad \text{(B)}$$

To avoid negative score values, we defined the final, rescaled EP risk score (*s*).

$$s = 0, \quad \text{if } 1.5 \cdot s_u + 18.95 < 0 \\ s = 15, \quad \text{if } 1.5 \cdot s_u + 18.95 > 15 \\ s = 1.5 \cdot s_u + 18.95, \quad \text{otherwise} \quad \text{(C)}$$

The EP risk score ranges from 0 to 15; higher values indicate a higher risk of recurrence.

EPclin (s_{clin}), a combined score consisting of the EP risk score and clinical parameters, was constructed from the training set:

$$s_{clin} = 0.35 \cdot t + 0.64 \cdot n + 0.28 \cdot s \quad \text{(D)}$$

where *t* codes the tumor size (1: $\leq$1 cm, 2: >1 to $\leq$2 cm, 3: >2 to $\leq$5 cm, and 4: >5 cm) and *n* the nodal status

(1: negative, 2: 1–3 positive nodes, 3: 4–10 positive nodes, and 4: >10 positive nodes).

Thresholds for EP and EPclin to discriminate patients into low and high risk of distant recurrence were developed in the training set at 5 and 3.3, respectively, and prespecified accordingly for validation.

Patients and tumor samples of the validation cohorts

The present investigation is part of the ABCSG translational research program (abcsg.research). Women included in the validation sets had participated in ABCSG-6 (tamoxifen-only arm) or ABCSG-8 trial and received tamoxifen for either 5 or 2 years followed by anastrozole for 3 years. Inclusion criteria and the main clinical results of these trials were reported previously (8, 9). Breast cancer patients with ER-negative and/or HER2-positive tumors were excluded from the analyses. Baseline clinical data of the validation cohorts are given in Supplementary Table S7. No samples from the validation cohorts were used for training. The clinical characteristics according to EP risk groups of the analyzed 378 patients of ABCSG-6 and the 1,324 patients of ABCSG-8 included in this study are listed in Supplementary Table S8. Of the successfully analyzed samples, 46 (12%, ABCSG-6) and 91 (7%, ABCSG-8) patients had distant recurrences at some point in time after surgery; the mean follow-up time was 97.4 (ABCSG-6) and 72.3 (ABCSG-8) months.

ABCSG-6 was a randomized phase III trial comparing tamoxifen alone for 5 years with tamoxifen in combination with aminoglutethimide for the first 2 years of treatment in postmenopausal women. In ABCSG-8, postmenopausal breast cancer patients were randomly assigned to receive tamoxifen for either 5 or 2 years followed by anastrozole for 3 years.

All FFPE tumor blocks were collected at the time of surgery prior to adjuvant therapy and were stored at room temperature. Approval was obtained from Institutional Review Boards. Tumor sections of 4 to 10 μm were cut. One section was stained by hematoxylin/eosin to confirm the presence of invasive carcinoma, and further sections were used for molecular analyses. Pathologists from participating ABCSG centers sent a representative FFPE tumor block from each woman to the central laboratory of abcsg.research at the Medical University of Vienna. Overall, 24 centers contributed samples (see Acknowledgments).

HER2 and Ki67 were assayed by IHC and evaluated centrally in the abcsg.research laboratory at the Medical University of Vienna according to standard criteria (10, 11). For Ki67 samples, a cutoff value of 11% was used for dichotomization into low or high expression (12). In case of HER2 2+ scores, FISH analyses were used to determine *ERBB2* gene amplification.

RNA extraction and gene expression analysis with quantitative reverse transcriptase PCR

Total RNA was extracted from one 5- or 10-μm whole FFPE tissue section with a silica bead–based, fully
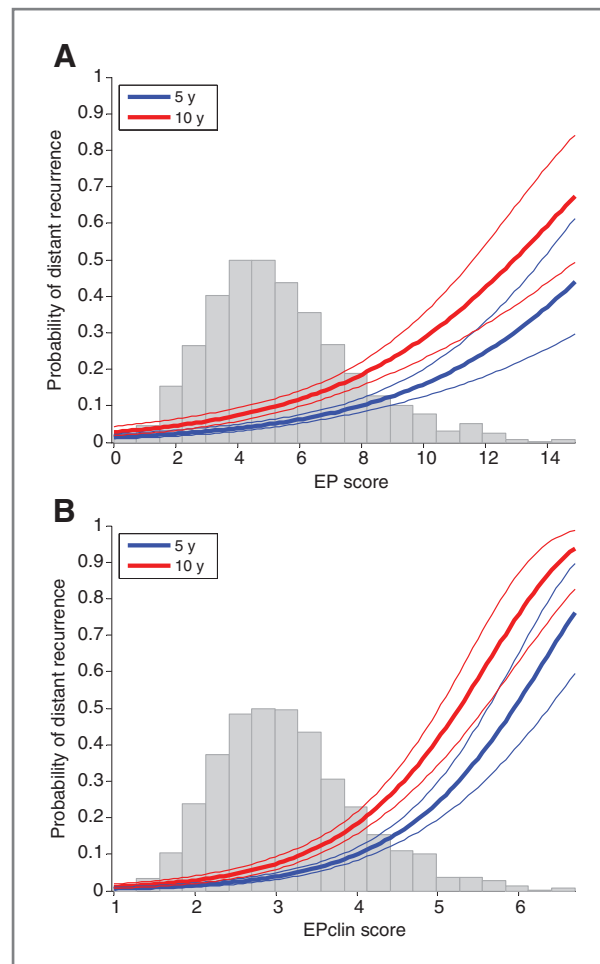

Figure 1. Estimated probability of distant recurrence as continuous functions of the EP risk score (A) and the EPclin risk score (B). The continuous relation between the respective score and the probability of developing a distant recurrence within the first 5 and 10 years after surgery is described by an independent model for each score generated from all ABCSG-6 and ABCSG-8 data ($n = 1,702$). The model is based on fitting the logarithm of the baseline cumulative hazard function by a cubic spline being a function of the logarithm of (event) time as proposed elsewhere (17). For both scores independently, the model describing the data best was found as having 2 degrees of freedom for the splines, the score as linear covariate, and no interaction between time and score. The thin curves indicate the 95% CI. The gray histogram in the background shows the distribution of scores for the patients.

automated isolation method (Tissue Preparation System, VERSANT Tissue Preparation Reagents, Siemens Healthcare Diagnostics) as described in Supplementary Appendix (13). All samples were analyzed with quantitative one-step reverse transcriptase PCR (qRT-PCR) on an ABI PRISM 7900HT (Applied Biosystems) using SuperScript III PLATINUM One-Step Quantitative RT-PCR System with ROX (Invitrogen). All PCR assays were conducted in triplicate (for details, see Supplementary Appendix). Normalized expression of GOI, EP, and EPclin was calculated as described above (Equations A–D).

Table 1. Multivariate Cox proportional hazard models for estimating the contribution of variables to predict distant recurrence

Cox model	Variable	ABCSG-6		ABCSG-8	
		Unit HR (95% CI)	*P*	Unit HR (95% CI)	*P*
Multivariate Cox model	Age	1.00 (0.96–1.04)	0.864	1.02 (0.99–1.04)	0.194
	Tumor size	1.09 (0.70–1.71)	0.704	1.57 (1.15–2.16)	0.005
	Nodal status	2.47 (1.75–3.48)	<0.001	2.32 (1.69–3.20)	<0.001
	Grade	0.81 (0.48–1.37)	0.435	1.09 (0.60–1.99)	0.770
	ER (IHC)	0.90 (0.58–1.40)	0.650	0.97 (0.70–1.34)	0.868
	PR (IHC)	0.83 (0.63–1.10)	0.199	0.94 (0.77–1.15)	0.559
	Ki67	1.03 (1.00–1.06)	0.086	1.00 (0.98–1.02)	0.974
	Treatment arm	–	–	0.78 (0.51–1.19)	0.243
	EP score	1.19 (1.04–1.36)	0.010	1.26 (1.15–1.38)	<0.001
Bivariate Cox model	Adjuvant! score	1.03 (1.02–1.04)	<0.001	1.05 (1.04–1.07)	<0.001
	EP score	1.19 (1.06–1.32)	0.002	1.27 (1.18–1.37)	<0.001

NOTE: The table shows 4 different Cox regression models with different sets of variables in both validation cohorts. Each model is shown in 2 columns: column "Unit HR" contains the 95% CIs for the unit HR of the respective variable whereas column "*P*" contains the *P* value for the probability of the regression coefficient to be 0. Variables/units were coded according to Supplementary Table S6 in the Supplementary Appendix.

Clinicopathologic risk classification

Adjuvant!Online (software version 8.0; www.adjuvantonline.com) was used to calculate the clinicopathologic risk based on patient's age, tumor size, tumor grade, ER status, and number of positive lymph nodes. Patients were considered as having low clinical risk when the 10-year risk as calculated by Adjuvant!Online was below 9% (14).

Validation and statistical analyses

The 2 validation studies were conducted using pre-specified objectives, assay methods, calculation procedures of scores, and cutoff values. Sample processing, qRT-PCR analyses, and calculations of scores were conducted completely blinded to clinical and outcome data by the laboratory scientists in Cologne. The scores and risk groups for each patient were then transferred to the ABCSG for statistical analysis in Vienna. The primary endpoint of the statistical analysis was distant recurrence. Statistical analysis of the validation study as well as coding of variables used in Cox regression, *c*-index calculations, and associated *P* values is described in more detail in section 3 of the Supplementary Appendix. Recurrence rates were estimated using the Kaplan–Meier method. All reported *P* values are results of 2-sided tests unless stated differently. Values of *P* < 5% were considered statistically significant.

As primary measure of the prognostic performance, the *c*-index was used (The Design Library: C-index; http://lib.stat.cmu.edu/S/Harrell/Design.html; ref. 15). Its relation to the time-dependent area under the receiver operator characteristic curve is shown in Supplementary Figure S8 in Supplementary Appendix. The *c*-index estimate

of a set of variables (e.g., the combination of the Adjuvant!Online score and the EP score) was calculated as described using 10,000 randomizations (15).

P values to assess whether the EP score contains additional information on distant recurrence with respect to a fixed set of variables X are 1-sided and based on the permutation test using 10,000 random permutations with null hypothesis "EP is not related to outcome" and the *c*-index as test statistic.

Statistical analyses were conducted using SPSS software, version 15.0 (SPSS Inc.) and MATLAB software, version R2009b (The MathWorks, Inc.).

Results

Validation of EP and EPclin in the ABCSG-6 and ABCSG-8 cohorts

The EP score was developed using gene expression data of 964 ER-positive, HER2-negative tumors (Supplementary Fig. S4, Supplementary Appendix, Section 2). Using the prespecified threshold for discriminating samples into low or high risk of distant metastasis, the EP risk score was independently validated in 2 large randomized phase III trials. As in the training set, all women had ER-positive, HER2-negative tumors and received adjuvant tamoxifen or tamoxifen–anastrozole treatment (Supplementary Fig. S7).

qRT-PCR was successful in 378 of 395 (95.7%) and 1,324 of 1,330 (99.5%) tumors of the 2 validation cohorts. The EP risk score as continuous predictor estimates the risk of distant recurrence in individual patients at 5 or 10 years (Fig. 1A).

The relation between the EP risk score and distant recurrence was assessed by Cox models adjusted for age,
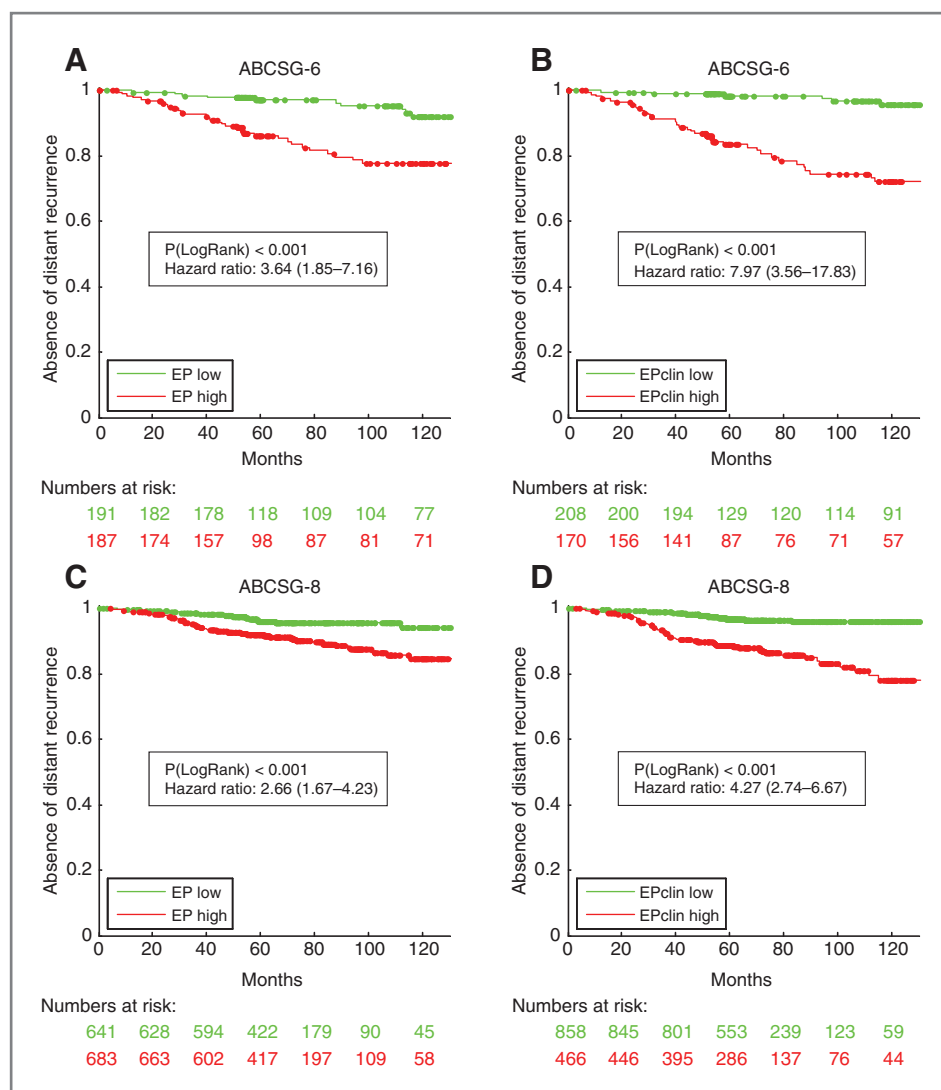
Figure 2. Kaplan–Meier plot of distant recurrence by EP and EPclin risk groups. Distant recurrence according to EP risk groups (A and C) and EPclin risk groups (B and D) in patients from the 2 validation cohorts (ABCSG-6 and ABCSG-8). Cutoff points for EP were prespecified at 5 (3.3 for EPclin) in the training set. Numbers in parentheses indicate the 95% CI of the HR.

tumor size, lymph node status, tumor grade, ER, progesterone receptor (PR), Ki67 expression (IHC), and type of adjuvant endocrine therapy. In these multivariate analyses, the EP risk score was an independent predictor of distant recurrence in ABCSG-6 and ABCSG-8 (Table 1).

Subsequently, patients were dichotomized into EP low- and EP high-risk groups according to the cutoff value prespecified in the training set.

In Kaplan–Meier analyses, we observed significant differences in distant recurrence between EP low-risk and EP high-risk patients in both trials (Fig. 2A and C). At 10 years, the distant recurrence rates for patients with EP low and EP high were 8% (3%–13%) and 22% (15%–29%) in ABCSG-6 ($P < 0.001$) and 6% (2%–9%) and 15% (11%–20%) in ABCSG-8 ($P < 0.001$), respectively.

In addition, the prespecified EPclin score was also validated in the ABCSG-6 and ABCSG-8 cohorts. EPclin is a continuous predictor of distant recurrence at 5 or 10 years (Fig. 1B). Distant recurrence rates of patients with
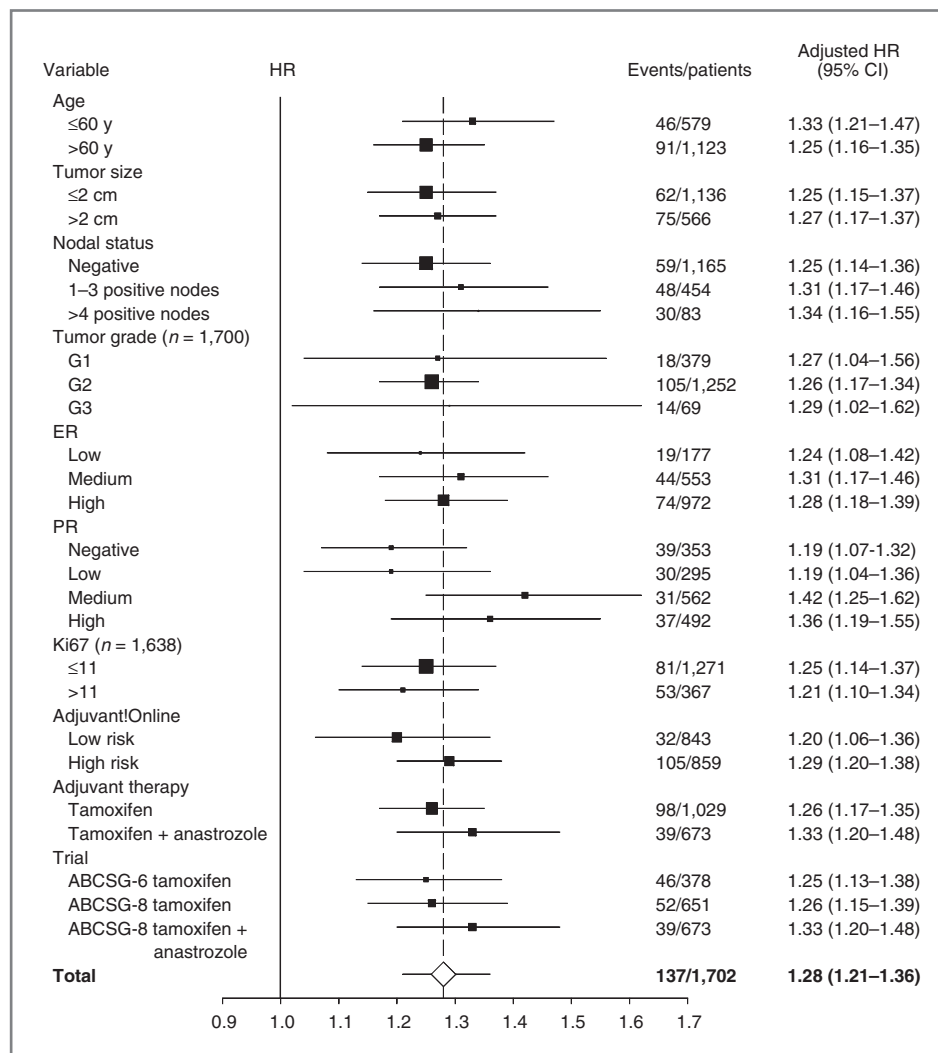
EPclin low and EPclin high were 4% (1%–8%) and 28% (20%–36%) in ABCSG-6 ($P < 0.001$) and 4% (2%–5%) and 22% (15%–29%) in ABCSG-8 ($P < 0.001$), respectively (Fig. 2B and D).

Performance of EP in subgroups

In subgroup analyses, we assessed potential heterogeneities of the prognostic performance of the continuous EP risk score by Cox regression analyses adjusted for the clinical risk as calculated by Adjuvant!Online. As shown in Figure 3, a higher EP risk score was associated with a significantly increased risk of distant recurrence in all analyzed subgroups including tumor size, nodal status, grading, and Ki67 and ER expression. In addition, there was no heterogeneity among the tamoxifen-only arm of ABCSG-6 and both treatment arms of ABCSG-8.

Exploratory subgroup analyses were conducted using Kaplan–Meier analyses with dichotomized EP low- and EP high-risk groups. We found significant differences in

Figure 3. Forest plot of the adjusted EP risk score HRs. The forest plot shows the adjusted EP risk score (as continuous variable) unit HRs (squares) and 95% CIs (horizontal lines) according to several clinical variables in the combined ABCSG-6 and ABCSG-8 cohort. HRs were adjusted for Adjuvant!Online. Sizes of the squares are proportional to the number of patients.

Variable	HR	Events/patients	Adjusted HR (95% CI)
Age			
≤60 y		46/579	1.33 (1.21–1.47)
>60 y		91/1,123	1.25 (1.16–1.35)
Tumor size			
≤2 cm		62/1,136	1.25 (1.15–1.37)
>2 cm		75/566	1.27 (1.17–1.37)
Nodal status			
Negative		59/1,165	1.25 (1.14–1.36)
1–3 positive nodes		48/454	1.31 (1.17–1.46)
>4 positive nodes		30/83	1.34 (1.16–1.55)
Tumor grade (n = 1,700)			
G1		18/379	1.27 (1.04–1.56)
G2		105/1,252	1.26 (1.17–1.34)
G3		14/69	1.29 (1.02–1.62)
ER			
Low		19/177	1.24 (1.08–1.42)
Medium		44/553	1.31 (1.17–1.46)
High		74/972	1.28 (1.18–1.39)
PR			
Negative		39/353	1.19 (1.07–1.32)
Low		30/295	1.19 (1.04–1.36)
Medium		31/562	1.42 (1.25–1.62)
High		37/492	1.36 (1.19–1.55)
Ki67 (n = 1,638)			
≤11		81/1,271	1.25 (1.14–1.37)
>11		53/367	1.21 (1.10–1.34)
Adjuvant!Online			
Low risk		32/843	1.20 (1.06–1.36)
High risk		105/859	1.29 (1.20–1.38)
Adjuvant therapy			
Tamoxifen		98/1,029	1.26 (1.17–1.35)
Tamoxifen + anastrozole		39/673	1.33 (1.20–1.48)
Trial			
ABCSG-6 tamoxifen		46/378	1.25 (1.13–1.38)
ABCSG-8 tamoxifen		52/651	1.26 (1.15–1.39)
ABCSG-8 tamoxifen + anastrozole		39/673	1.33 (1.20–1.48)
Total		**137/1,702**	**1.28 (1.21–1.36)**

0.9 1.0 1.1 1.2 1.3 1.4 1.5 1.6 1.7

distant recurrence between EP low-risk and EP high-risk patients in small (T1) and large (T2/T3) tumors, in node-negative and node-positive patients, in grade 1 and 2, in ER low (≤80%) and high (>80%), in Ki67 low (≤11%, luminal A) tumors as well as in Adjuvant!Online low (<9%) and high (≥9%) patients (Supplementary Fig. S9).

EP score adds statistically significant information to established clinicopathologic risk factors

To assess the contribution of different clinicopathologic and molecular variables to the prediction of distant recurrence, we calculated unbiased estimates of the c-index for several combinations of variables (Fig. 4). Combination of the established prognostic markers [nodal status, tumor size, age, grade, quantitative ER (IHC), Ki67, and treatment (ABCSG-8 only)] resulted in c-indices of 0.705 in ABCSG-6 and 0.700 in ABCSG-8. However, addition of the molecular EP score to these clinicopathologic markers resulted in a significant improvement of the prognostic power

(c-indices: 0.727 in ABCSG-6 and 0.728 in ABCSG-8). Similar results were obtained when using the Adjuvant!Online score as predefined multicomponent clinicopathologic risk assessment: c-indices significantly increased from 0.749 to 0.785 in ABCSG-6 and from 0.690 to 0.733 in ABCSG-8 by addition of the molecular EP score. Because quantitative ER (IHC, within ER-positive samples) and Ki67 are not included in the Adjuvant!Online score, we examined the combination of these variables with Adjuvant!Online and assessed the additional prognostic information of EP. We could show that EP significantly added prognostic power also to these clinicopathologic variables [Adjuvant!Online, quantitative ER, Ki67, and treatment (ABCSG-8 only)] in both cohorts (0.750 vs. 0.764 in ABCSG-6 and 0.701 vs. 0.726 in ABCSG-8). Most interestingly, the EPclin score had higher c-indices than all combinations of clinicopathologic variables reported above (0.788 in ABCSG-6 and 0.732 in ABCSG-8).
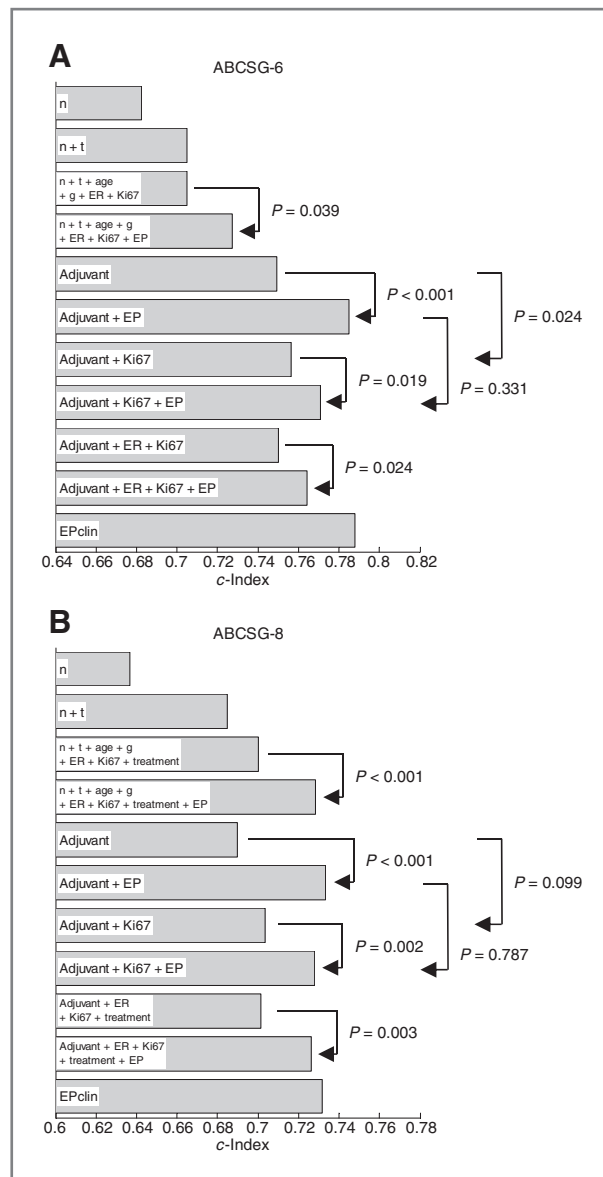
Figure 4. *c*-Index indicating the performance of combinations of different prognostic factors in the ABCSG-6 (A) and -8 (B) studies. n, nodal status; t, tumor size; g, grade; ER, estrogen receptor; EP, Endopredict. The values on the *x*-axis are unbiased estimates of the c-index of the linear combination of one or more variables by Cox regression. Statistical tests indicate whether the *c*-index increases significantly by addition of EP to a fixed set of clinicopathologic variables (see section 3.3 of the Supplementary Appendix).

Figure 5. Comparison of the Adjuvant!Online score and EPclin (ABCSG-6 and ABCSG-8). The subgroups of samples concordant or discordant between Adjuvant!Online and EPclin are shown. Cutoff points were 9% for Adjuvant!Online and 3.3 for EPclin.

Discussion

In this study, we developed and validated a new RNA-based molecular test which predicts distant recurrence in ER-positive, HER2-negative early-stage breast cancer patients. The EP score significantly adds prognostic information to established clinicopathologic factors such as Adjuvant!Online in combination with quantitative ER expression and Ki67. This is a differentiating feature not shown by other breast cancer prognosis tests so far. Moreover, we generated and validated a combined molecular/clinicopathologic score (EPclin) which outperformed the conventional risk factors alone. EPclin might be helpful in therapeutic decision making when the use of therapy in addition to endocrine therapy is uncertain after consideration of conventional markers. With our results, we reach level I evidence according to Simon and colleagues because our study consisted of 2 category B studies using archived samples from 2 similar but separate prospective clinical trials (16).

The EP risk score has several advantages compared with currently available or published prognostic multigene assays. Identification of the candidate genes was based on 3 predefined requirements. First, we used a large discovery cohort with tumor samples of nearly 1,000 patients. Second, we analyzed a homogeneous breast cancer subgroup consisting of ER-positive, HER2-negative patients only. Finally, we exclusively included patients treated homogeneously with adjuvant tamoxifen in the training set. Using this approach, we minimized confounding effects from other tumor subtypes or anticancer drugs unrelated to the primary research question. Furthermore, we have chosen a conservative approach for algorithm

Finally, in an exploratory analysis, we investigated the distant recurrence rate of patients with tumors classified discordantly between Adjuvant!Online and EPclin. We found 467 (27.4%) samples with discordant classification: Adjuvant!Online high risk and EPclin low risk ($n = 345$) or Adjuvant!Online low risk and EPclin high risk ($n = 122$). Within the subgroup of discordant samples outcome follows the EPclin classification rather than the Adjuvant!Online classification ($P = 0.0496$; Fig. 5).

design aiming at selecting the most robust assembly of prognostic genes, which resulted in a set of 8 GOI and 3 reference genes, all measured in triplicate. The requirement of both, platform (HG-U133A arrays and qRT-PCR) and sample type (fresh frozen and FFPE) transferability, assessed in a large cohort of matched tumor samples is expected to contribute to assay robustness, an important feature when the test will be used in clinical routine. Moreover, for the performance of the test, no special preparation, storage, or shipment of tissue is necessary, and assessment of gene expression was possible in 99% of the samples. This is why the test is expected to be compatible to a standard clinical workflow of breast cancer patient care using easily available FFPE tissue.

The EP risk score has not only prognostic value independent of established clinicopathologic risk factors in each of the 2 validation studies, ABCSG-6 and ABCSG-8, but also provides significant additional prognostic information to Adjuvant!Online score, quantitative ER IHC, and Ki67 labeling index. The combination of the molecular EP risk score with the clinicopathologic risk factors nodal status and tumor size resulted in the EPclin score identifying risk groups with striking differences in the 10-year distant recurrence rates. EPclin low-risk patients had a 10-year risk of distant recurrence of 4% and had, therefore, adequately been treated with adjuvant endocrine therapy only. In contrast, for EPclin high-risk patients with a 10-year distant recurrence risk of 28% (ABCSG-6) and 22% (ABCSG-8) adjuvant endocrine therapy alone may not be sufficient and additional or other adjuvant treatment may be indicated. A limitation of our study is that the optimal adjuvant therapy in the EPclin high-risk patient population remains unknown for now and needs to be determined in well-designed prospective clinical trials. A further limitation is that the validation studies have a lower number of grade 3 tumors than what would be expected in a population of postmenopausal breast cancer. However, EP-based risk classification was of borderline significance in this small subgroup of patients.

In summary, using FFPE tumor tissue, the EP risk score significantly predicted distant recurrence of breast cancer patients. The EP risk score provided prognostic information independent of—and in addition to—clinicopathologic variables, in particular to Adjuvant!Online and Ki67 labeling index. By combining the EP risk score with clinicopathologic risk factors, we identified breast cancer patients with low risk for distant recurrence who may be adequately and sufficiently treated with adjuvant endocrine therapy only. Using this new easy-to-perform multigene tool in clinical practice has a strong potential to assist in optimizing adjuvant therapy by reducing both undertreatment and overtreatment and thus might

improve management of patients with early-stage breast cancer.

Disclosure of Potential Conflicts of Interest

M. Filipits received honoraria from Speakers Bureau of Astra Zeneca, Eli Lilly, and Sividon Diagnostics GmbH and is a consultant/advisory board member of Astra Zeneca. P. Dubsky received honoraria from Speakers Bureau of Novartis and Pfizer and is a consultant/advisory board member of Novartis. C.F. Singer received honoraria from Speakers Bureau of Novartis and Roche. H. Brauch received honoraria from Speakers Bureau of Astra Zeneca. M. Schwab received honoraria from Speakers Bureau of Roche Molecular Diagnostics. K.E. Weber has ownership interest (including patents) from Sividon Diagnostics GmbH. R. Kronenwett has ownership interest (including patents) from Sividon Diagnostics GmbH. M. Gnant received honoraria from Speakers Bureau of Astra Zeneca, Novartis, and Sanofi-Aventis and is a consultant/advisory board member of Amgen, Astra Zeneca, Novartis, Roche, and Sanofi-Aventis. The other authors disclosed no potential conflicts of interest.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

References

1. Goldhirsch A, Ingle JN, Gelber RD, Coates AS, Thurlimann B, Senn HJ. Thresholds for therapies: highlights of the St Gallen International Expert Consensus on the primary therapy of early breast cancer 2009. Ann Oncol 2009;20:1319–29.

2. Carlson RW, Allred DC, Anderson BO, Burstein HJ, Carter WB, Edge SB, et al. Breast cancer. Clinical practice guidelines in oncology. J Natl Compr Canc Netw 2009;7: 122–92.

3. Sotiriou C, Pusztai L. Gene-expression signatures in breast cancer. N Engl J Med 2009;360:790–800.

4. van de Vijver MJ, He YD, van't Veer LJ, Dai H, Hart AA, Voskuil DW, et al. A gene-expression signature as a predictor of survival in breast cancer. N Engl J Med 2002;347:1999–2009.

5. Wang Y, Klijn JG, Zhang Y, Sieuwerts AM, Look MP, Yang F, et al. Gene-expression profiles to predict distant metastasis of lymph-node-negative primary breast cancer. Lancet 2005;365: 671–9.

6. Paik S, Shak S, Tang G, Kim C, Baker J, Cronin M, et al. A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer. N Engl J Med 2004;351:2817–26.

7. Albain KS, Barlow WE, Shak S, Hortobagyi GN, Livingston RB, Yeh IT, et al. Prognostic and predictive value of the 21-gene recurrence score assay in postmenopausal women with node-positive, oestrogen-receptor-positive breast cancer on chemo-therapy: a retrospective analysis of a randomised trial. Lancet Oncol 2009;11:55–65.

8. Schmid M, Jakesz R, Samonigg H, Kubista E, Gnant M, Menzel C, et al. Randomized trial of tamoxifen versus tamoxifen plus aminoglu-tethimide as adjuvant treatment in postmenopausal breast cancer patients with hormone receptor-positive disease: Austrian breast and colorectal cancer study group trial 6. J Clin Oncol 2003;21:984–90.

9. Jakesz R, Jonat W, Gnant M, Mittlboeck M, Greil R, Tausch C, et al. Switching of postmenopausal women with endocrine-responsive early breast cancer to anastrozole after 2 years' adjuvant tamoxifen: combined results of ABCSG trial 8 and ARNO 95 trial. Lancet 2005;366:455–62.

10. Pohl G, Rudas M, Taucher S, Stranzl T, Steger GG, Jakesz R, et al. Expression of cell cycle regulatory proteins in breast carcinomas before and after preoperative chemotherapy. Breast Cancer Res Treat 2003;78:97–103.

11. Reiner-Concin A, Regitnig P, Dinges HP, Hofler G, Lax S, Muller-Holzner E, et al. Practice of HER-2 immunohistochemistry in breast carcinoma in Austria. Pathol Oncol Res 2008;14:253–9.

12. Viale G, Giobbie-Hurder A, Regan MM, Coates AS, Mastropasqua MG, Dell'Orto P, et al. Prognostic and predictive value of centrally reviewed Ki-67 labeling index in postmenopausal women with endo-crine-responsive breast cancer: results from Breast International Group Trial 1-98 comparing adjuvant tamoxifen with letrozole. J Clin Oncol 2008;26:5569–75.

13. Hennig G, Gehrmann M, Stropp U, Brauch H, Fritz P, Eichelbaum M, et al. Automated extraction of DNA and RNA from a single formalin-fixed paraffin-embedded tissue section for analysis of both single-nucleotide polymorphisms and mRNA expression. Clin Chem 2010;56:1845–53.

14. Buyse M, Loi S, van't Veer L, Viale G, Delorenzi M, Glas AM, et al. Validation and clinical utility of a 70-gene prognostic signature for women with node-negative breast cancer. J Natl Cancer Inst 2006; 98:1183–92.

15. Parker JS, Mullins M, Cheang MC, Leung S, Voduc D, Vickery T, et al. Supervised risk predictor of breast cancer based on intrinsic sub-types. J Clin Oncol 2009;27:1160–7.

16. Simon RM, Paik S, Hayes DF. Use of archived specimens in evaluation of prognostic and predictive biomarkers. J Natl Cancer Inst 2009; 101:1446–52.

17. Royston P, Parmar MK. Flexible parametric proportional-hazards and proportional-odds models for censored survival data, with application to prognostic modelling and estimation of treatment effects. Stat Med 2002;21:2175–97.